

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Koji YANAI et al.
- (ii) TITLE OF INVENTION: β -FRUCTOFURANOSIDASE AND ITS GENE,
METHOD OF ISOLATING β -FRUCTOFURANOSIDASE GENE, SYSTEM
FOR PRODUCING β -FRUCTOFURANOSIDASE, AND
 β -FRUCTOFURANOSIDASE VARIANT
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
 - (B) STREET: 2033 K Street, N.W., Suite 800
 - (C) CITY: Washington
 - (D) STATE: D.C.
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 20006
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: MS-DOS
 - (D) SOFTWARE: Wordperfect 5.1
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: NEW
 - (B) FILING DATE: November 23, 2001
- (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/142,623

(B) FILING DATE: September 10, 1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Lee Cheng

(B) REGISTRATION NUMBER: 40,949

(C) REFERENCE/DOCKET NUMBER: 2001-1611

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 202-721-8200

(B) TELEFAX: 202-721-8250

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 635 amino acid residues

(B) TYPE: amino acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1

(ATCC 20611)

(ix) FEATURE:

(A) NAME/KEY: mat peptide

(B) LOCATION: 1..635

(C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ser Tyr His Leu Asp Thr Thr Ala Pro Pro Pro Thr Asn Leu Ser Thr
1 5 10 15
Leu Pro Asn Asn Thr Leu Phe His Val Trp Arg Pro Arg Ala His Ile
20 25 30
Leu Pro Ala Glu Gly Gln Ile Gly Asp Pro Cys Ala His Tyr Thr Asp
35 40 45
Pro Ser Thr Gly Leu Phe His Val Gly Phe Leu His Asp Gly Asp Gly
50 55 60
Ile Ala Gly Ala Thr Thr Ala Asn Leu Ala Thr Tyr Thr Asp Thr Ser
65 70 75 80
Asp Asn Gly Ser Phe Leu Ile Gln Pro Gly Gly Lys Asn Asp Pro Val
85 90 95
Ala Val Phe Asp Gly Ala Val Ile Pro Val Gly Val Asn Asn Thr Pro
100 105 110
Thr Leu Leu Tyr Thr Ser Val Ser Phe Leu Pro Ile His Trp Ser Ile
115 120 125
Pro Tyr Thr Arg Gly Ser Glu Thr Gln Ser Leu Ala Val Ala Arg Asp
130 135 140
Gly Gly Arg Arg Phe Asp Lys Leu Asp Gln Gly Pro Val Ile Ala Asp
145 150 155 160
His Pro Phe Ala Val Asp Val Thr Ala Phe Arg Asp Pro Phe Val Phe
165 170 175
Arg Ser Ala Lys Leu Asp Val Leu Leu Ser Leu Asp Glu Glu Val Ala
180 185 190

Arg	Asn	Glu	Thr	Ala	Val	Gln	Gln	Ala	Val	Asp	Gly	Trp	Thr	Glu	Lys				
195					200					205									
Asn	Ala	Pro	Trp	Tyr	Val	Ala	Val	Ser	Gly	Gly	Val	His	Gly	Val	Gly				
210					215					220									
Pro	Ala	Gln	Phe	Leu	Tyr	Arg	Gln	Asn	Gly	Gly	Asn	Ala	Ser	Glu	Phe				
225					230					235					240				
Gln	Tyr	Trp	Glu	Tyr	Leu	Gly	Glu	Trp	Trp	Gln	Glu	Ala	Thr	Asn	Ser				
245					250					255									
Ser	Trp	Gly	Asp	Glu	Gly	Thr	Trp	Ala	Gly	Arg	Trp	Gly	Phe	Asn	Phe				
260					265					270									
Glu	Thr	Gly	Asn	Val	Leu	Phe	Leu	Thr	Glu	Glu	Gly	His	Asp	Pro	Gln				
275					280					285									
Thr	Gly	Glu	Val	Phe	Val	Thr	Leu	Gly	Thr	Glu	Gly	Ser	Gly	Leu	Pro				
290					295					300									
Ile	Val	Pro	Gln	Val	Ser	Ser	Ile	His	Asp	Met	Leu	Trp	Ala	Ala	Gly				
305					310					315					320				
Glu	Val	Gly	Val	Gly	Ser	Glu	Gln	Glu	Gly	Ala	Lys	Val	Glu	Phe	Ser				
325					330					335									
Pro	Ser	Met	Ala	Gly	Phe	Leu	Asp	Trp	Gly	Phe	Ser	Ala	Tyr	Ala	Ala				
340					345					350									
Ala	Gly	Lys	Val	Leu	Pro	Ala	Ser	Ser	Ala	Val	Ser	Lys	Thr	Ser	Gly				
355					360					365									
Val	Glu	Val	Asp	Arg	Tyr	Val	Ser	Phe	Val	Trp	Leu	Thr	Gly	Asp	Gln				
370					375					380									

Tyr	Glu	Gln	Ala	Asp	Gly	Phe	Pro	Thr	Ala	Gln	Gln	Gly	Trp	Thr	Gly
385					390					395					400
Ser	Leu	Leu	Leu	Pro	Arg	Glu	Leu	Lys	Val	Gln	Thr	Val	Glu	Asn	Val
				405					410					415	
Val	Asp	Asn	Glu	Leu	Val	Arg	Glu	Glu	Gly	Val	Ser	Trp	Val	Val	Gly
			420					425					430		
Glu	Ser	Asp	Asn	Gln	Thr	Ala	Arg	Leu	Arg	Thr	Leu	Gly	Ile	Thr	Ile
		435					440					445			
Ala	Arg	Glu	Thr	Lys	Ala	Ala	Leu	Leu	Ala	Asn	Gly	Ser	Val	Thr	Ala
	450					455					460				
Glu	Glu	Asp	Arg	Thr	Leu	Gln	Thr	Ala	Ala	Val	Val	Pro	Phe	Ala	Gln
465					470					475				480	
Ser	Pro	Ser	Ser	Lys	Phe	Phe	Val	Leu	Thr	Ala	Gln	Leu	Glu	Phe	Pro
				485					490				495		
Ala	Ser	Ala	Arg	Ser	Ser	Pro	Leu	Gln	Ser	Gly	Phe	Glu	Ile	Leu	Ala
		500					505					510			
Ser	Glu	Leu	Glu	Arg	Thr	Ala	Ile	Tyr	Tyr	Gln	Phe	Ser	Asn	Glu	Ser
		515					520				525				
Leu	Val	Val	Asp	Arg	Ser	Gln	Thr	Ser	Ala	Ala	Ala	Pro	Thr	Asn	Pro
		530				535					540				
Gly	Leu	Asp	Ser	Phe	Thr	Glu	Ser	Gly	Lys	Leu	Arg	Leu	Phe	Asp	Val
545				550					555					560	
Ile	Glu	Asn	Gly	Gln	Glu	Gln	Val	Glu	Thr	Leu	Asp	Leu	Thr	Val	Val
			565						570				575		

Val Asp Asn Ala Val Val Glu Val Tyr Ala Asn Gly Arg Phe Ala Leu		
580	585	590
Ser Thr Trp Ala Arg Ser Trp Tyr Asp Asn Ser Thr Gln Ile Arg Phe		
595	600	605
Phe His Asn Gly Glu Gly Glu Val Gln Phe Arg Asn Val Ser Val Ser		
610	615	620
Glu Gly Leu Tyr Asn Ala Trp Pro Glu Arg Asn		
625	630	635

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1905 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double stranded
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1
(ATCC 20611)

(ix) FEATURE:

- (A) NAME/KEY: mat peptide
- (B) LOCATION: 1 .. 1905
- (C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCATACCACC	TGGACACCAC	GGCCCCGCCG	CCGACCAACC	TCAGCACCCCT	CCCCAACAAC	60
ACCCTCTTCC	ACGTGTGGCG	GCCGCGCGCG	CACATCCTGC	CCGCCGAGGG	CCAGATCGGC	120
GACCCCTGCG	CGCACTACAC	CGACCCATCC	ACCGGCCTCT	TCCACGTGGG	G TTCCTGCAC	180
GACGGGGACG	GCATCGCGGG	CGCCACCACG	GCCAACCTGG	CCACCTACAC	CGATACCTCC	240
GATAACGGGA	GCTTCCTGAT	CCAGCCGGGC	GGGAAGAACG	ACCCCGTCGC	CGTGTTTCGAC	300
GGCGCCGTCA	TCCCCGTGCG	CGTCAACAAC	ACCCCCACCT	TACTCTACAC	CTCCGTCTCC	360
TTCCTGCCCA	TCCACTGGTC	CATCCCCTAC	ACCCGCGGCA	GCGAGACGCA	GTCGTTGGCC	420
GTCGCGCGCG	ACGGCGGCCG	CCGCTTCGAC	AAGCTCGACC	AGGGCCCCGT	CATCGCCGAC	480
CACCCCTTCG	CCGTCGACGT	CACCGCCTTC	CGCGATCCGT	TTGTCTTCCG	CAGTGCCAAG	540
TTGGATGTGC	TGCTGTCGTT	GGATGAGGAG	GTGGCGCGGA	ATGAGACGGC	CGTGCAGCAG	600
GCCGTCGATG	GCTGGACCGA	GAAGAACGCC	CCCTGGTATG	TCGCGGTCTC	TGGCGGGGTG	660
CACGGCGTCG	GGCCCGCGCA	G TTCCTCTAC	CGCCAGAACG	GCGGGAACGC	TTC CGAGTTC	720
CAGTACTGGG	AGTACCTCGG	GGAGTGGTGG	CAGGAGGCGA	CCA ACTCCAG	CTGGGGCGAC	780
GAGGGCACCT	GGGCCGGGCG	CTGGGGGTTC	AACTTCGAGA	CGGGGAATGT	GCTCTTCCTC	840
ACCGAGGAGG	GCCATGACCC	CCAGACGGGC	GAGGTGTTTC	TCACCCTCGG	CACGGAGGGG	900
TCTGGCCTGC	CAATCGTGCC	GCAGGTCTCC	AGTATCCACG	ATATGCTGTG	GGCGCGGGT	960
GAGGTCGGGG	TGGGCAGTGA	GCAGGAGGGT	GCCAAGGTCG	AGTTCTCCCC	CTCCATGGCC	1020
GGGTTTCTGG	ACTGGGGGTT	CAGCGCCTAC	GCTGCGGCGG	GCAAGGTGCT	GCCGGCCAGC	1080
TCGGCGGTGT	CGAAGACCAG	CGGCGTGGAG	GTGGATCGGT	ATGTCTCGTT	CGTCTGGTTG	1140
ACGGGCGACC	AGTACGAGCA	GGCGGACGGG	TTCCCCACGG	CCCAGCAGGG	GTGGACGGGG	1200
TCGCTGCTGC	TGCCGCGCGA	GCTGAAGGTG	CAGACGGTGG	AGAACGTCGT	CGACAACGAG	1260
CTGGTGCGCG	AGGAGGGCGT	GTCGTGGGTG	GTGGGGGAGT	CGGACAACCA	GACGGCCAGG	1320
CTGCGCACGC	TGGGGATCAC	GATCGCCCGG	GAGACCAAGG	CGGCCCTGCT	GGCCAACGGC	1380
TCGGTGACCG	CGGAGGAGGA	CCGCACGCTG	CAGACGGCGG	CCGTCTGTCC	G TTCGCGCAA	1440
TCGCCGAGCT	CCAAGTTCTT	CGTGCTGACG	GCCCAGCTGG	AGTTCCCCGC	GAGCGCGCGC	1500

TCGTCCCCGC TCCAGTCCGG GTTCGAAATC CTGGCGTCGG AGCTGGAGCG CACGGCCATC 1560
TACTACCAGT TCAGCAACGA GTCGCTGGTC GTCGACCGCA GCCAGACTAG TGCGGCGGGCG 1620
CCCACGAACC CCGGGCTGGA TAGCTTTACT GAGTCCGGCA AGTTGCGGTT GTTCGACGTG 1680
ATCGAGAACG GCCAGGAGCA GGTCGAGACG TTGGATCTCA CTGTCGTCGT GGATAACGCG 1740
GTTGTGCGAGG TGTATGCCAA CGGGCGCTTT GCGTTGAGCA CCTGGGCGAG ATCGTGGTAC 1800
GACAACTCCA CCCAGATCCG CTTCTTCCAC AACGGCGAGG GCGAGGTGCA GTTCAGGAAT 1860
GTCTCCGTGT CGGAGGGGCT CTATAACGCC TGGCCGGAGA GAAAT 1905

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acid residues
- (B) TYPE: Amino acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Peptide

(v) FRAGMENT TYPE: internal fragment

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1
(ATCC 20611)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Leu Asp Gln Gly Pro Val Ile Ala Asp His Pro Phe Ala Val Asp Val
1 5 10 15
Thr Ala Phe Arg
20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acid residues

(B) TYPE: Amino acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Peptide

(v) FRAGMENT TYPE: internal fragment

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1
(ATCC 20611)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Val Glu Phe Ser Pro Ser Met Ala Gly Phe Leu Asp Trp Gly Phe Ser

1 5 10 15

Ala Tyr Ala Ala

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acid residues

(B) TYPE: Amino acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Peptide

(v) FRAGMENT TYPE: internal fragment

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
(ATCC 20611)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Val Gln Thr Val Glu Asn Val Val Asp Asn Glu Leu Val Arg Glu Glu
1 5 10 15
Gly Val Ser Trp
20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acid residues
(B) TYPE: Amino acid
(C) STRANDEDNESS: Not relevant
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Peptide

(v) FRAGMENT TYPE: internal fragment

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: Aspergillus niger ACE-2-1
(ATCC 20611)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ala Ala Leu Leu Ala Xaa Gly Ser Val Thr Ala Glu Glu Asp Arg Thr
1 5 10 15
Leu Gln Thr Ala
20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acid residues
- (B) TYPE: Amino acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Peptide

(v) FRAGMENT TYPE: N-terminal fragment

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1
(ATCC 20611)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ser Tyr His Leu Asp Thr

1

5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATCGCSGAYC AYCCSTTYGC 20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCRTRTCSA CSACRTTYTC 20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 788 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double stranded
- (D) TOPOLOGY: Linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Microorganism: *Aspergillus niger* ACE-2-1
(ATCC 20611)

(ix) FEATURE:

- (A) NAME/KEY: P CDS(partial amino acid sequence)
- (B) LOCATION: 1 .. 788
- (C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATC GCC GAC CAC CCC TTC GCC GTC GAC GTC ACC GCC TTC CGC GAT CCG	48
Ile Ala Asp His Pro Phe Ala Val Asp Val Thr Ala Phe Arg Asp Pro	
1 5 10 15	
TTT GTC TTC CGC AGT GCC AAG TTG GAT GTG CTG CTG TCG TTG GAT GAG	96
Phe Val Phe Arg Ser Ala Lys Leu Asp Val Leu Leu Ser Leu Asp Glu	
20 25 30	
GAG GTG GCG CGG AAT GAG ACG GCC GTG CAG CAG GCC GTC GAT GGC TGG	144
Glu Val Ala Arg Asn Glu Thr Ala Val Gln Gln Ala Val Asp Gly Trp	
35 40 45	
ACC GAG AAG AAC GCC CCC TGG TAT GTC GCG GTC TCT GGC GGG GTG CAC	192
Thr Glu Lys Asn Ala Pro Trp Tyr Val Ala Val Ser Gly Gly Val His	
50 55 60	
GGC GTC GGG CCC GCG CAG TTC CTC TAC CGC CAG AAC GGC GGG AAC GCT	240
Gly Val Gly Pro Ala Gln Phe Leu Tyr Arg Gln Asn Gly Gly Asn Ala	
65 70 75 80	
TCC GAG TTC CAG TAC TGG GAG TAC CTC GGG GAG TGG TGG CAG GAG GCG	288
Ser Glu Phe Gln Tyr Trp Glu Tyr Leu Gly Glu Trp Trp Gln Glu Ala	
85 90 95	
ACC AAC TCC AGC TGG GGC GAC GAG GGC ACC TGG GCC GGG CGC TGG GGG	336
Thr Asn Ser Ser Trp Gly Asp Glu Gly Thr Trp Ala Gly Arg Trp Gly	
100 105 110	
TTC AAC TTC GAG ACG GGG AAT GTG CTC TTC CTC ACC GAG GAG GGC CAT	384
Phe Asn Phe Glu Thr Gly Asn Val Leu Phe Leu Thr Glu Glu Gly His	
115 120 125	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

GAC CCC CAG ACG GGC GAG GTG TTC GTC ACC CTC GGC ACG GAG GGG TCT	432
Asp Pro Gln Thr Gly Glu Val Phe Val Thr Leu Gly Thr Glu Gly Ser	
130 135 140	
GGC CTG CCA ATC GTG CCG CAG GTC TCC AGT ATC CAC GAT ATG CTG TGG	480
Gly Leu Pro Ile Val Pro Gln Val Ser Ser Ile His Asp Met Leu Trp	
145 150 155 160	
GCG GCG GGT GAG GTC GGG GTG GGC AGT GAG CAG GAG GGT GCC AAG GTC	528
Ala Ala Gly Glu Val Gly Val Gly Ser Glu Gln Glu Gly Ala Lys Val	
165 170 175	
GAG TTC TCC CCC TCC ATG GCC GGG TTT CTG GAC TGG GGG TTC AGC GCC	576
Glu Phe Ser Pro Ser Met Ala Gly Phe Leu Asp Trp Gly Phe Ser Ala	
180 185 190	
TAC GCT GCG GCG GGC AAG GTG CTG CCG GCC AGC TCG GCG GTG TCG AAG	624
Tyr Ala Ala Ala Gly Lys Val Leu Pro Ala Ser Ser Ala Val Ser Lys	
195 200 205	
ACC AGC GGC GTG GAG GTG GAT CGG TAT GTC TCG TTC GTC TGG TTG ACG	672
Thr Ser Gly Val Glu Val Asp Arg Tyr Val Ser Phe Val Trp Leu Thr	
210 215 220	
GGC GAC CAG TAC GAG CAG GCG GAC GGG TTC CCC ACG GCC CAG CAG GGG	720
Gly Asp Gln Tyr Glu Gln Ala Asp Gly Phe Pro Thr Ala Gln Gln Gly	
225 230 235 240	
TGG ACG GGG TCG CTG CTG CTG CCG CGC GAG CTG AAG GTG CAG ACG GTG	768
Trp Thr Gly Ser Leu Leu Leu Pro Arg Glu Leu Lys Val Gln Thr Val	
245 250 255	

GAG AAC GTC GTC GAC AAC GA

788

Glu Asn Val Val Asp Asn

260

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 565 amino acid residues

(B) TYPE: Amino acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: *Penicillium roqueforti* IAM7254

(ix) FEATURE:

(A) NAME/KEY: mat peptide

(B) LOCATION: 1 .. 565

(C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Asp Phe His Thr Pro Ile Asp Tyr Asn Ser Ala Pro Pro Asn Leu

1 5 10 15

Ser Thr Leu Ala Asn Ala Ser Leu Phe Lys Thr Trp Arg Pro Arg Ala

20 25 30

His Leu Leu Pro Pro Ser Gly Asn Ile Gly Asp Pro Cys Gly His Tyr

35 40 45

Thr	Asp	Pro	Lys	Thr	Gly	Leu	Phe	His	Val	Gly	Trp	Leu	Tyr	Ser	Gly
50						55						60			
Ile	Ser	Gly	Ala	Thr	Thr	Asp	Asp	Leu	Val	Thr	Tyr	Lys	Asp	Leu	Asn
65					70					75					80
Pro	Asp	Gly	Ala	Pro	Ser	Ile	Val	Ala	Gly	Gly	Lys	Asn	Asp	Pro	Leu
				85					90					95	
Ser	Val	Phe	Asp	Gly	Ser	Val	Ile	Pro	Ser	Gly	Ile	Asp	Gly	Met	Pro
				100				105					110		
Thr	Leu	Leu	Tyr	Thr	Ser	Val	Ser	Tyr	Leu	Pro	Ile	His	Trp	Ser	Ile
		115						120					125		
Pro	Tyr	Thr	Arg	Gly	Ser	Glu	Thr	Gln	Ser	Leu	Ala	Val	Ser	Tyr	Asp
	130					135					140				
Gly	Gly	His	Asn	Phe	Thr	Lys	Leu	Asn	Gln	Gly	Pro	Val	Ile	Pro	Thr
145					150					155					160
Pro	Pro	Phe	Ala	Leu	Asn	Val	Thr	Ala	Phe	Arg	Asp	Pro	Tyr	Val	Phe
				165					170					175	
Gln	Ser	Pro	Ile	Leu	Asp	Lys	Ser	Val	Asn	Ser	Thr	Gln	Gly	Thr	Trp
		180						185					190		
Tyr	Val	Ala	Ile	Ser	Gly	Gly	Val	His	Gly	Val	Gly	Pro	Cys	Gln	Phe
		195						200					205		
Leu	Tyr	Arg	Gln	Asn	Asp	Ala	Asp	Phe	Gln	Tyr	Trp	Glu	Tyr	Leu	Gly
	210					215					220				
Gln	Trp	Trp	Lys	Glu	Pro	Leu	Asn	Thr	Thr	Trp	Gly	Lys	Gly	Asp	Trp
225					230					235					240

Ala Gly Gly Trp Gly Phe Asn Phe Glu Val Gly Asn Val Phe Ser Leu			
245	250	255	
Asn Ala Glu Gly Tyr Ser Glu Asp Gly Glu Ile Phe Ile Thr Leu Gly			
260	265	270	
Ala Glu Gly Ser Gly Leu Pro Ile Val Pro Gln Val Ser Ser Ile Arg			
275	280	285	
Asp Met Leu Trp Val Thr Gly Asn Val Thr Asn Asp Gly Ser Val Thr			
290	295	300	
Phe Lys Pro Thr Met Ala Gly Val Leu Asp Trp Gly Val Ser Ala Tyr			
305	310	315	320
Ala Ala Ala Gly Lys Ile Leu Pro Ala Ser Ser Gln Ala Ser Thr Lys			
325	330	335	
Ser Gly Ala Pro Asp Arg Phe Ile Ser Tyr Val Trp Leu Thr Gly Asp			
340	345	350	
Leu Phe Glu Gln Val Lys Gly Phe Pro Thr Ala Gln Gln Asn Trp Thr			
355	360	365	
Gly Ala Leu Leu Leu Pro Arg Glu Leu Asn Val Arg Thr Ile Ser Asn			
370	375	380	
Val Val Asp Asn Glu Leu Ser Arg Glu Ser Leu Thr Ser Trp Arg Val			
385	390	395	400
Ala Arg Glu Asp Ser Gly Gln Ile Asp Leu Glu Thr Met Gly Ile Ser			
405	410	415	
Ile Ser Arg Glu Thr Tyr Ser Ala Leu Thr Ser Gly Ser Ser Phe Val			
420	425	430	

Glu	Ser	Gly	Lys	Thr	Leu	Ser	Asn	Ala	Gly	Ala	Val	Pro	Phe	Asn	Thr
435				440				445							
Ser	Pro	Ser	Ser	Lys	Phe	Phe	Val	Leu	Thr	Ala	Asn	Ile	Ser	Phe	Pro
450				455				460							
Thr	Ser	Ala	Arg	Asp	Ser	Gly	Ile	Gln	Ala	Gly	Phe	Gln	Val	Leu	Ser
465				470				475				480			
Ser	Ser	Leu	Glu	Ser	Thr	Thr	Ile	Tyr	Tyr	Gln	Phe	Ser	Asn	Glu	Ser
485				490				495							
Ile	Ile	Val	Asp	Arg	Ser	Asn	Thr	Ser	Ala	Ala	Ala	Arg	Thr	Thr	Ala
500				505				510							
Gly	Ile	Leu	Ser	Asp	Asn	Glu	Ala	Gly	Arg	Leu	Arg	Leu	Phe	Asp	Val
515				520				525							
Leu	Arg	Asn	Gly	Lys	Glu	Gln	Val	Glu	Thr	Leu	Glu	Leu	Thr	Ile	Val
530				535				540							
Val	Asp	Asn	Ser	Val	Leu	Glu	Val	Tyr	Ala	Asn	Gly	Arg	Phe	Ala	Leu
545				550				555				560			
Gly	Thr	Trp	Ala	Arg											
565															

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1695 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double stranded
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: *Penicillium roqueforti* IAM7254

(ix) FEATURE:

(A) NAME/KEY: mat peptide

(B) LOCATION: 1 .. 1695

(C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTTGATTTCC ATACCCCGAT TGACTATAAC TCGGCTCCGC CAAACCTTTC TACCCTGGCA	60
AACGCATCTC TTTTCAAGAC ATGGAGACCC AGAGCCCATC TTCTCCCTCC ATCTGGGAAC	120
ATAGGCGACC CGTGCGGGCA CTATACCGAT CCCAAGACTG GTCTCTTCCA CGTGGGTTGG	180
CTTTACAGTG GGATTTCTGGG AGCGACAACC GACGATCTCG TTACCTATAA AGACCTCAAT	240
CCCGATGGAG CCCCGTCAAT TGTTCAGGA GGAAAGAACG ACCCTCTTTC TGTCTTCGAT	300
GGCTCGGTCA TTCCAAGCGG TATAGACGGC ATGCCAACTC TTCTGTATAC CTCTGTATCA	360
TACCTCCCAA TCCACTGGTC CATCCCCTAC ACCCGGGGAA GCGAGACACA ATCCTTGGCC	420
GTTTCCTATG ACGGTGGTCA CAACTTCACC AAGCTCAACC AAGGGCCCGT GATCCCTACG	480
CCTCCGTTTG CTCTCAATGT CACCGCTTTC CGTGACCCCT ACGTTTTCCTA AAGCCCAATT	540
CTGGACAAAT CTGTCAATAG TACCCAAGGA ACATGGTATG TCGCCATATC TGGCGGTGTC	600
CACGGTGTCTG GACCTTGTCA GTTCCTCTAC CGTCAGAACG ACGCAGATTT TCAATATTGG	660
GAATATCTCG GGCAATGGTG GAAGGAGCCC CTTAATACCA CTTGGGGAAA GGGTGAAGTG	720
GCCGGGGGTT GGGGCTTCAA CTTTGAGGTT GGCAACGTCT TTAGTCTGAA TGCAGAGGGG	780
TATAGTGAAG ACGGCGAGAT ATTCATAACC CTCGGTGCTG AGGGTTCGGG ACTTCCCATC	840
GTTCTCTAAG TCTCCTCTAT TCGCGATATG CTGTGGGTGA CCGGCAATGT CACAAATGAC	900
GGCTCTGTCA CTTTCAAGCC AACCATGGCG GGTGTGCTTG ACTGGGGCGT GTCGGCATAT	960
GCTGCTGCAG GCAAGATCTT GCCGGCCAGC TCTCAGGCAT CCACAAAGAG CGGTGCCCCC	1020

GATCGGTTCA	TTTCCTATGT	CTGGCTCACT	GGAGATCTAT	TCGAGCAAGT	GAAAGGATTC	1080
CCTACCGCTC	AACAAAAC TG	GACCGGGGCC	CTCTTACTGC	CGCGAGAGCT	GAATGTCCGC	1140
ACTATCTCTA	ACGTGGTGGA	TAACGAACTT	TCGCGTGAGT	CCTTGACATC	GTGGCGCGTG	1200
GCCCCGGAAG	ACTCTGGTCA	GATCGACCTT	GAAACAATGG	GAATCTCAAT	TTCCAGGGAG	1260
ACTTACAGCG	CTCTCACATC	CGGCTCATCT	TTTGTGAGT	CTGGTAAAAC	GTTGTCGAAT	1320
GCTGGAGCAG	TGCCCTTCAA	TACCTCACCC	TCAAGCAAGT	TCTTCGTGCT	GACAGCAAAT	1380
ATATCTTTCC	CGACCTCTGC	CCGTGACTCT	GGCATCCAGG	CTGGTTTCCA	GGTTTTATCC	1440
TCTAGTCTTG	AGTCTACAAC	TATCTACTAC	CAATTCTCCA	ACGAGTCCAT	CATCGTCGAC	1500
CGCAGCAACA	CGAGTGCTGC	GGCGAGAACA	ACTGCTGGGA	TCCTCAGTGA	TAACGAGGCG	1560
GGACGTCTGC	GCCTCTTCGA	CGTGTTGCGA	AATGGAAAAG	AACAGGTTGA	AACTTTGGAG	1620
CTCACTATCG	TGGTGGATAA	TAGTGTACTG	GAAGTATATG	CCAATGGACG	CTTTGCTCTA	1680
GGCACTTGGG	CTCGG					1695

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 574 amino acid residues
- (B) TYPE: Amino acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Microorganism: *Scopulariopsis brevicaulis* IF04843

(ix) FEATURE:

- (A) NAME/KEY: mat peptide
- (B) LOCATION: 1 .. 574

(C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gln Pro Thr Ser Leu Ser Ile Asp Asn Ser Thr Tyr Pro Ser Ile Asp
1 5 10 15
Tyr Asn Ser Ala Pro Pro Asn Leu Ser Thr Leu Ala Asn Asn Ser Leu
20 25 30
Phe Glu Thr Trp Arg Pro Arg Ala His Val Leu Pro Pro Gln Asn Gln
35 40 45
Ile Gly Asp Pro Cys Met His Tyr Thr Asp Pro Glu Thr Gly Ile Phe
50 55 60
His Val Gly Trp Leu Tyr Asn Gly Asn Gly Ala Ser Gly Ala Thr Thr
65 70 75 80
Glu Asp Leu Val Thr Tyr Gln Asp Leu Asn Pro Asp Gly Ala Gln Met
85 90 95
Ile Leu Pro Gly Gly Val Asn Asp Pro Ile Ala Val Phe Asp Gly Ala
100 105 110
Val Ile Pro Ser Gly Ile Asp Gly Lys Pro Thr Met Met Tyr Thr Ser
115 120 125
Val Ser Tyr Met Pro Ile Ser Trp Ser Ile Ala Tyr Thr Arg Gly Ser
130 135 140
Glu Thr His Ser Leu Ala Val Ser Ser Asp Gly Gly Lys Asn Phe Thr
145 150 155 160
Lys Leu Val Gln Gly Pro Val Ile Pro Ser Pro Pro Phe Gly Ala Asn
165 170 175

Val Thr Ser Trp Arg Asp Pro Phe Leu Phe Gln Asn Pro Gln Phe Asp			
180	185	190	
Ser Leu Leu Glu Ser Glu Asn Gly Thr Trp Tyr Thr Val Ile Ser Gly			
195	200	205	
Gly Ile His Gly Asp Gly Pro Ser Ala Phe Leu Tyr Arg Gln His Asp			
210	215	220	
Pro Asp Phe Gln Tyr Trp Glu Tyr Leu Gly Pro Trp Trp Asn Glu Glu			
225	230	235	240
Gly Asn Ser Thr Trp Gly Ser Gly Asp Trp Ala Gly Arg Trp Gly Tyr			
245	250	255	
Asn Phe Glu Val Ile Asn Ile Val Gly Leu Asp Asp Asp Gly Tyr Asn			
260	265	270	
Pro Asp Gly Glu Ile Phe Ala Thr Val Gly Thr Glu Trp Ser Phe Asp			
275	280	285	
Pro Ile Lys Pro Gln Ala Ser Asp Asn Arg Glu Met Leu Trp Ala Ala			
290	295	300	
Gly Asn Met Thr Leu Glu Asp Gly Asp Ile Lys Phe Thr Pro Ser Met			
305	310	315	320
Ala Gly Tyr Leu Asp Trp Gly Leu Ser Ala Tyr Ala Ala Ala Gly Lys			
325	330	335	
Glu Leu Pro Ala Ser Ser Lys Pro Ser Gln Lys Ser Gly Ala Pro Asp			
340	345	350	
Arg Phe Val Ser Tyr Leu Trp Leu Thr Gly Asp Tyr Phe Glu Gly His			
355	360	365	

Asp Phe Pro Thr Pro Gln Gln Asn Trp Thr Gly Ser Leu Leu Leu Pro
 370 375 380
 Arg Glu Leu Ser Val Gly Thr Ile Pro Asn Val Val Asp Asn Glu Leu
 385 390 395 400
 Ala Arg Glu Thr Gly Ser Trp Arg Val Gly Thr Asn Asp Thr Gly Val
 405 410 415
 Leu Glu Leu Val Thr Leu Lys Gln Glu Ile Ala Arg Glu Thr Leu Ala
 420 425 430
 Glu Met Thr Ser Gly Asn Ser Phe Thr Glu Ala Ser Arg Asn Val Ser
 435 440 445
 Ser Pro Gly Ser Thr Ala Phe Gln Gln Ser Leu Asp Ser Lys Phe Phe
 450 455 460
 Val Leu Thr Ala Ser Leu Ser Phe Pro Ser Ser Ala Arg Asp Ser Asp
 465 470 475 480
 Leu Lys Ala Gly Phe Glu Ile Leu Ser Ser Glu Phe Glu Ser Thr Thr
 485 490 495
 Val Tyr Tyr Gln Phe Ser Asn Glu Ser Ile Ile Ile Asp Arg Ser Asn
 500 505 510
 Ser Ser Ala Ala Ala Leu Thr Thr Asp Gly Ile Asp Thr Arg Asn Glu
 515 520 525
 Phe Gly Lys Met Arg Leu Phe Asp Val Val Glu Gly Asp Gln Glu Arg
 530 535 540
 Ile Glu Thr Leu Asp Leu Thr Ile Val Val Asp Asn Ser Ile Val Glu
 545 550 555 560

Val His Ala Asn Gly Arg Phe Ala Leu Ser Thr Trp Val Arg

565

570

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1722 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double stranded

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Microorganism: *Scopulariopsis brevicaulis* IFO4843

(ix) FEATURE:

(A) NAME/KEY: mat peptide

(B) LOCATION: 1 .. 1722

(C) IDENTIFICATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CAACCTACGT CTCTGTCAAT CGACAATTCC ACGTATCCTT CTATCGACTA CAACTCCGCC	60
CCTCCAAACC TCTCGACTCT TGCCAACAAC AGCCTCTTCG AGACATGGAG GCCGAGGGCA	120
CACGTCCTTC CGCCCCAGAA CCAGATCGGC GATCCGTGTA TGCACTACAC CGACCCCGAG	180
ACAGGAATCT TCCACGTCGG CTGGCTGTAC AACGGCAATG GCGCTTCCGG CGCCACGACC	240
GAGGATCTCG TCACCTATCA GGATCTCAAC CCCGACGGAG CGCAGATGAT CCTTCCGGGT	300
GGTGTGAATG ACCCCATTGC TGTCTTTGAC GGC GCGGTTA TTCCCAGTGG CATTGATGGG	360
AAACCCACCA TGATGTATAC CTCGGTGTCA TACATGCCCA TCTCCTGGAG CATCGCTTAC	420
ACCAGGGGAA GCGAGACCCA CTCTCTCGCA GTGTCGTCCG ACGGCGGTAA GAACTTCACC	480

AAGCTGGTGC	AGGGCCCCGT	CATTCCTTCG	CCTCCCTTCG	GCGCCAACGT	GACCAGCTGG	540
CGTGACCCCT	TCCTGTTCCA	AAACCCCCAG	TTCGACTCTC	TCCTCGAAAG	CGAGAACGGC	600
ACGTGGTACA	CCGTTATCTC	TGGTGGCATC	CACGGTGACG	GCCCCTCCGC	GTTCTCTAC	660
CGTCAGCACG	ACCCCGACTT	CCAGTACTGG	GAGTACCTTG	GACCGTGGTG	GAACGAGGAA	720
GGGAACTCGA	CCTGGGGCAG	CGGTGACTGG	GCTGGCCGGT	GGGGCTACAA	CTTCGAGGTC	780
ATCAACATTG	TCGGTCTTGA	CGATGATGGC	TACAACCCCG	ACGGTGAAAT	CTTTGCCACG	840
GTAGGTACCG	AATGGTCGTT	TGACCCCATC	AAACCGCAGG	CCTCGGACAA	CAGGGAGATG	900
CTCTGGGCCG	CGGGCAACAT	GACTCTCGAG	GACGGCGATA	TCAAGTTCAC	GCCAAGCATG	960
GCGGGCTACC	TCGACTGGGG	TCTATCGGCG	TATGCCGCCG	CTGGCAAGGA	GCTGCCCGCT	1020
TCTTCAAAGC	CTTCGCAGAA	GAGCGGTGCG	CCGGACCGGT	TCGTGTCGTA	CCTGTGGCTC	1080
ACCGGTGACT	ACTTCGAGGG	CCACGACTTC	CCCACCCCGC	AGCAGAATTG	GACCGGCTCG	1140
CTTTTGCTTC	CGCGTGAGCT	GAGCGTCGGG	ACGATTCCCA	ACGTTGTCGA	CAACGAGCTT	1200
GCTCGCGAGA	CGGGCTCTTG	GAGGGTTGGC	ACCAACGACA	CTGGCGTGCT	TGAGCTGGTC	1260
ACTCTGAAGC	AGGAGATTGC	TCGCGAGACG	CTGGCTGAAA	TGACCAGCGG	CAACTCCTTC	1320
ACCGAGGCGA	GCAGGAATGT	CAGCTCGCCC	GGATCTACCG	CCTTCCAGCA	GTCCCTGGAT	1380
TCCAAGTTCT	TCGTCCTGAC	CGCCTCGCTC	TCCTTCCCTT	CGTCGGCTCG	CGACTCCGAC	1440
CTCAAGGCTG	GTTTCGAGAT	CCTGTCGTCC	GAGTTTGAGT	CGACCACGGT	CTACTACCAG	1500
TTTTCCAACG	AGTCCATCAT	CATTGACCGG	AGCAACTCGA	GTGCTGCCGC	CTTGACTACC	1560
GATGGAATCG	ACACCCGCAA	CGAGTTTGGC	AAGATGCGCC	TGTTTGATGT	TGTCGAGGGT	1620
GACCAGGAGC	GTATCGAGAC	GCTCGATCTC	ACTATTGTGG	TTGATAACTC	GATCGTTGAG	1680
GTTTCATGCCA	ACGGGCGATT	CGCTCTGAGC	ACTTGGGTTC	GG		1722

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGAATTCCA ATGAAGCTCA CCACTACC

28

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCGGATCCCG GTCAATTTCT CTCC

24

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GACTGACCGG TGTTTCATCC

19

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTCGGTTGTC ATAGATGTGG

20

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CAATCCAGGA GGATCCCAAT GAAG

24

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Synthetic DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGACCGGGAT CCGGGCATGC AG

22

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Not relevant
 - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Synthetic DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGCGTCGTCT AGAGGTTGTC ACTT

24

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: Nucleic acid
 - (C) STRANDEDNESS: Not relevant
 - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCCTATTGGG GTCCATGGCC C

21

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CAACTGCTGG CATCCTCAGT GA

22

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GCGGATCCAT GAAGCTATCA AATGCAATCA

30

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCGGATCCTT ACCGAGCCCA AGTGCC

26

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCGGATCCAA TGAAGCTCAC CACTACC

27

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCGGATCCCG GTCAATTTCT CTCC

24

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCACCGCCT GGCGCGATCC G

21

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGCACGGAGT GGTCTGGCC

19

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTCCAGTATC AAGGATATGC TGTC

24

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGACCAGTAC AAGCAGGCGG

20

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TCCAGTATCC GCGATATGCT G

21

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CGGCACGGAG GTTCTGGCC TGC

23

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Not relevant

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CGGCACGGAG GAGTCTGGCC TGC

23

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Not relevant
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Synthetic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CGGCACGGAG GATTCTGGCC TGC

23